The Valuation of Interest Rate Derivatives: Empirical Evidence from the Spanish Market.[‡]

Juan M. Moraleda * Ton C.F. Vorst **

October 28, 1996

Abstract

This paper studies empirical issues of one-factor yield curve models. We focus on the models by Ho & Lee (1986), Hull & White (1990) and Moraleda & Vorst (1996). To be consistent in the comparison of the models, we derive them all within the Ritkchen and Sankarasubramanian (1995) framework, which is a subset of the very general Heath, Jarrow and Morton (1992) model. We estimate model parameters from historical time series of government bond prices. The model by Moraleda and Vorst (1996) turns out to best explain the yield curve dynamics through time. Moreover, humped shapes in the volatility structure as modelled in this model are typically found. Next, we use these parameter estimations for pricing options traded in the Spanish financial market. A comparison between model and market option prices is provided.

1 Introduction.

Over the last two decades, a lot of research has been devoted to the development of valuation models for interest rate derivative securities such as options, caps, collars and swaptions. Most attention has been paid to the theoretical properties of these models, where very few papers have studied empirical issues. Most of these empirical works have focussed

[‡]A version of this paper under the title "Empirical behavior of interest rate models" is to appear in "System Dynamics in Economic and Financial Models", Ed Wiley, 1997. A previous version of this paper was entitled "Pricing Derivative Securities with One Factor Yield Curve Models".

*J.M. Moraleda Novo, Erasmus University Rotterdam and Universidad Autonoma de Madrid. Tinbergen Institute, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands, Tel: +31.10.408.14.65, Fax: +31.10.452.73.47, E-mail: moraleda@tir.few.eur.nl

**T.C.F. Vorst, Department of Finance, Erasmus University Rotterdam and Erasmus Center for Financial Research, P.O. Box 1738, 300 Rotterdam, The Netherlands, Tel: +31.10.408.12.70, Fax: +31.10.452.77.46, E-mail: vorst@opres.few.eur.nl

We would like to thank Henk Hoek and Fabio Mercurio for helpful comments.

on what the econometrians would call diagnostic checking or specification testing of the models. This means checking the assumptions of the model, computing statistics of the parameter estimators, studying their stability over time, etc... Typically, this strand of literature tests the models with bond price data. Brown and Dybvig (1986), Stambaugh (1988), Dybvig (1989), Chan, Karolyi, Longstaff and Sanders (1992) and Brown and Schaefer (1994) are examples of research where the main purpose is estimating different model parameters and subsequently testing with bond price data.

The main result from these studies is that relatively simple models such as Vasicek (1977) or Cox, Ingersoll and Rox (1985) fit bond price data remarkably well. A step further in the analysis of the empirical behavior of the term-structure models is, though, to test how the models with these parameter values price real option data. This is of crucial interest since these models are mainly used for derivative pricing purposes. In our opinion, this is the central question to be answered by empirical work in this area. Surprisingly, it has been rarely addressed in the literature. Some examples are Flesaker (1993), who analysed the continuous time version of the Ho and Lee (1986) model and Dietrich-Campbell and Schwartz (1986), who examined the two factor Brennan and Schwartz (1986) model.

Some literature has approached this central question estimating the model parameters with option data rather than with the underlying yield curve movements. In fact, the model parameters are estimated in the same way as the implied volatility for options on stocks is estimated from option and stock price data. Amin and Morton (1994) and Amin and Ng (1995) are probably the most relevant examples that estimate implied volatilities and comparing different models. However, this approach has some theoretical inconsistencies. In particular, implied volatilities techniques demand a different estimation on each trading day. The result is that the model parameters change daily. Contrary, term-structure models generally assume that the volatility parameters are constant over time. The quoted authors have tried to reconcile this apparent inconsistency by different arguments although a full understanding of the implied volatility estimation for models assuming constant volatility parameters has not been yet obtained[1].

In this paper, we focus on the estimation of term-structure model parameters from historical yield curve movements. As the previous literature, we analyze econometric issues such as the stability of the parameters. Our main interest, however, concerns the reliability of the estimated parameters to price options traded in the markets.

We use the general Heath, Jarrow and Morton (1992) framework for pricing options. However, for most interest rate derivative securities, such as American options, explicit calculation of prices in this kind of framework requires computationally intensive numerical methods. Recently, however, Li, Ritchken and Sankarasubramanian (1995) have developed a numerical algorithm that makes the evolution of the term structure Markovian for a broad class of models, as identified by Ritchken and Sankarasubramanian (1995). All models considered in this paper can be included in this class. We estimate parameters for the Spanish market by minimizing the error between the forward rate changes produced by the models and those historically observed in the markets. The forward rates, however,

[1]See Amin and Morton (1994)

are not directly observed in the markets. They rather have to be inferred from bond prices quoted in the debt markets as will be explained in Section 3. The parameters thus estimated are used to price derivative securities traded in the Spanish option markets. A comparison between the option prices given by the models and the prices quoted in the markets is provided.

This paper compares three one-factor yield curve models. They have one, two and three parameters respectively. The first one is the continuous time version of the Ho and Lee (1986) model. Ho and Lee are credited for being the first to build a model that provides arbitrage free prices which depend on an exogenously specified initial yield curve. Moreover, this process, contrary to traditional models, no longer involves an exogenous specification of the 'market price of risk'.

The Ho and Lee model, however, has the major disadvantage of just allowing parallel yield curves for any moment in the future. This is because the volatility structure of the interest rate dynamics is assumed to be a constant.

The second model under consideration is the Hull and White (1990) model or, equivalently, the exponentially decaying model by Heath, Jarrow and Morton(1992). This model adds a mean-reverting effect for the interest rates through a second parameter in the volatility specification. In fact, the volatility structure under this model is a strictly decreasing function of the time to maturity so that short term interest rates are more volatile than long term rates. As a result, this models captures the well known effect that interest rates are pulled to some long-run average level over time.

Some recent empirical studies, however, have found that the mean reversion for the interest rates is not actually as straightforward as it was generally believed. In fact, the volatility structure is not, frequently, a monotone decreasing function of the time to maturity as modeled by Hull and White (1990). But, rather, it is initially upwards sloping, it reaches a maximum, and then, due to the mean reversion, it decreases with the time to maturity. This shape in the volatility for the interest rates is referred to in the financial literature as humped volatility structures.[2]

The third model we consider in this paper is due to Moraleda and Vorst (1996). This model allows for humped volatility structures in the yield curve dynamics at the price of adding an extra parameter to the model.

The rest of this paper is organized as follows. In section 2, the general setting in which all the models are nested is presented. Section 3 describes the two different data set that we will use. They record yield curve data and option price data, respectively. Section 4 outlines the methodology for the estimation of the models, while section 5 presents the results achieved on both the volatility estimation and the fit to the option price data. Section 6 draws the main conclusions and summarizes the paper.

[2] See, for example, Kahn (1991), Heath, Jarrow, Morton and Spindel (1992) and Amin and Morton (1994) that have found, for different periods and currencies, humped volatility structures in the yield curve dynamics.

2 The models.

The yield curve dynamics can be equivalently modeled through three related variables: interest rates, forward rates and prices of discount bonds. The traditional approach focusses on the interest rate modeling. Vasicek (1977), for example, proposed the following characterization of the instantaneous spot interest rate dynamics

$$dr = (a - br)dt + \sigma dW(t), \tag{1}$$

where a,b and σ are non-negative constants, r is short hand for $r(t)$, the spot interest rate, and $W(t)$ is a Brownian motion. Extensions of this stochastic differential equation yield to well known models in the financial literature.

More recently, Heath, Jarrow and Morton (1992) have proposed to model the forward rates rather than the interest rate dynamics. We derive all considered models under this framework as we outline in the sequel.

Consider a continuous time economy where bonds are traded for all maturities and markets are frictionless. Denote by $P(t,T)$ the price at time t of a pure discount bond that pays \$ 1 at time T, and assume that $P(t,T) > 0$ for all $t \in [0,T]$. The instantaneous forward rate at time t for a maturity T, $f(t,T)$, is defined by

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T},$$

so that

$$P(t,T) = e^{-\int_t^T f(t,u)du}. \tag{2}$$

For a fixed maturity T, Heath, Jarrow and Morton (1992) model the evolution of the instantaneous forward rates by the diffusion

$$df(t,T) = \alpha(t,T,f(t,T))dt + \sigma(t,T,f(t,T))dW(t), \tag{3}$$

with $f(0,T)$ given and deterministic, and where $\alpha(.)$ and $\sigma(.)$ are stochastic processes whose values are known at time t and $W(t)$ is a Brownian motion. Notice that both $\alpha(.)$ and $\sigma(.)$ can explicitly depend on the forward rates $f(t,T)$. The dependence of $\alpha(.)$ and $\sigma(.)$ on $f(t,T)$ will not be made explicit hereafter.

Heath, Jarrow and Morton (1992) proved that at any moment in time arbitrage free prices $g(t)$ of interest rates derivative securities with only a terminal payoff $g(T)$ at time T are given by

$$g(t) = E_t\left(e^{-\int_t^T r(u)du} g(T)\right), \tag{4}$$

where the expectation is taken with respect to the so called risk adjusted process which is specified by equation (3) with the restriction that

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,u)du. \tag{5}$$

Substituting (5) in (3) and integrating, gives the instantaneous forward rate process under which theoretical arbitrage free prices should be calculated

$$f(t,T) = f(0,T) + \int_0^t \sigma(u,T)\left(\int_u^T \sigma(u,y)dy\right)du + \int_0^t \sigma(u,T)d\tilde{W}(u), \tag{6}$$

where $\tilde{W}(u)$ denotes a Brownian motion under an equivalent martingale measure.

As follows from (6), the stochastic evolution of the forward rates under the risk neutral process is fully characterized with the specification of an initial forward rate curve and the volatility function. The initial forward rate curve is observable in the market at any moment in time. The volatility function, in turn, plays a key role in the analysis. Its specification uniquely determines the drift term of the risk adjusted process by the no arbitrage argument.

The spot rate at time t, $r(t)$, is given by

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma(u,t)\left(\int_u^t \sigma(u,y)dy\right)du + \int_0^t \sigma(u,t)d\tilde{W}(u). \tag{7}$$

Once this general setting has been established, we can now discuss the particular models under consideration. Consider the framework given by (6) with

$$\sigma(t,T) = \sigma\frac{[1+\gamma T]}{[1+\gamma t]}e^{-\frac{\lambda}{2}(T-t)} \tag{8}$$

and where σ, γ and λ are non-negative constants. This choice for the volatility generalizes a number of term-structure models. In particular, we consider the following models.

Table 1: One-Factor Yield Curve Models

Model	Key	Volatility specification
1. Ho and Lee (1986)	HL	$\sigma(t,T) = \sigma$
2. Hull and White (1990)	HW	$\sigma(t,T) = \sigma e^{-\frac{\lambda}{2}(T-t)}$
3. Moraleda and Vorst (1996)	MV	$\sigma(t,T) = \sigma\frac{[1+\gamma T]}{[1+\gamma t]}e^{-\frac{\lambda}{2}(T-t)}$

1. *The constant volatility model (HL).* If we set $\lambda = \gamma = 0$, we get the continuous time version of the Ho and Lee (1986) model as derived by Heath, Jarrow and Morton (1992). This model is a one-parameter model with the volatility function being constant. Hereafter, this model is referred to as the HL model.

2. *The exponential volatility model (HW).* This second model is the extended Vasicek version of the Hull and White (1990) model or, equivalently, the exponentially decaying model by Heath, Jarrow and Morton (1992). It is obtained by setting $\gamma = 0$,

which leads to Vasicek's volatility function[3]. This model generalizes the previous model by Ho and Lee (1986). Indeed, for $\lambda = 0$ in the HW model, we get the constant volatility model. What is more important, though, is that this approach models a very well known effect for the interest rates dynamics known as mean-reversion. This means that interest rates are pulled over time to some long-run average level. Hence, it seems to imply that short term rates are more volatile than long term rates. This model is denoted along this paper as the exponential volatility or HW model.

3. *The humped volatility model (MV).* We finally consider a three-parameters model introduced by Moraleda and Vorst (1996). These authors model yield curve dynamics with a volatility function as given by (8) with all parameters being strictly positive. This yields a humped volatility structure (for $2\gamma > \lambda$) that has been systematically found in recent empirical studies. Notice that this model still implies a mean-reverting process for the interest rates, as strongly supported by both economic theory and previous empirical evidence. What this model adds to the existing literature is a delay of the mean reverting effect. In fact, the volatility structure is upwards sloping for the short-term rates. After reaching a maximum, the volatility decreases with time so that interest rates revert to their long-run average. After all, the empirical findings by Kahn (1991), Amin and Morton (1994), etc... point out that the mean-reverting effect is not as straightforward as it was generally believed. And this is precisely what Moraleda and Vorst (1996) modelled[4]. This model shares all the advantages of the previous models and generalizes them. We refer to this model as the humped volatility or MV model.

As explained in the introduction, the final goal of this chapter is to test the models with real option data. If the options to be priced were European, their valuation would be fairly simple since equation (4) can be used. In fact, equation (4) can be explicitly further developed into an analytical formula in case the volatility is described by (8) (see Mercurio and Moraleda (1996)). Therefore, pricing European style claims just requires solving an

[3]Brenner (1989) proved that the function $\sigma \exp\left[-\frac{\lambda}{2}(T-t)\right]$ in (6) yields an Ornstein-Uhlenbeck process for the spot rates as assumed by Vasicek (1977). While the exponentially decaying model by Heath, Jarrow and Morton (1992) matches the initial term structure by construction, Hull and White (1990) allowed the drift parameter of Vasicek's model to be time dependent so that they exogenously fit the initial yield curve. As a result, both models are equivalent [see Moraleda and Vorst (1996) for a formal proof]. Though, Hull and White (1990) were the first to incorporate the initial yield curve as observed in the markets to a model with Vasicek's volatility. For this historical reason we choose their name to refer to this model. However, as with the HL model, the derivation of the model that we use through out this paper is due to Heath, Jarrow and Morton (1992).

[4]The Moraleda and Vorst (1996) model as given by (8) is somehow related to a previous model by Mercurio and Moraleda (1996) where humped volatilities are already modelled. The latter authors considered a volatility structure given by $\sigma(t,T) = \sigma[1+\gamma(T-t)]\exp[-\frac{\lambda}{2}(T-t)]$, which is humped (for $2\gamma > \lambda$) and stationary. But, such a choice leads to a model that cannot be used for pricing American-style claims with a recombining lattice. On the other hand, the Moraleda and Vorst (1996) model can price American options with recombining lattices, although it is no longer stationary.

analytic formula.[5]

For the case of American style options, things are not so easy. Since we use this kind of option data to test the models, we devote the next subsection to explain how to price these claims within a homogeneous framework for all models.

2.1 The valuation of American style claims

In contrast with the European options case, there are no analytic formulas for the valuation of American style claims. The expectation of the terminal payoff of the security considered should be calculated recursively down to the valuation date t. This means discretising the SDE (3) which leads, in the general case of general case of Heath, Jarrow and Morton (1992), to a non-Markov process for the short rate, r This makes computing the model very slow when not unfeasible. Indeed, for one factor models, there are 2^n nodes at the $n-th$ time step. Recently, however, Ritchken and Sankarasubramanian (1995) have identified a class of volatility structures within the Heath, Jarrow and Morton (1992) paradigm that enable the evolution of the term structure to be made Markovian with respect to two state variables and thus reduce the number of different nodes at each instant considerably. As we show in this section, this class is fairly general. In fact, we show here how to embed all models reported in Table 1 in the class of Ritchken and Sankarasubramanian (1995).

The Ritchken and Sankarasubramanian (1995) class of models is given by equation (3), where we have the following specification for the volatility

$$\begin{aligned} \sigma(t,T) &= \sigma(t)h(t,T), \\ h(t,T) &= e^{-\int_t^T \kappa(x)dx}, \end{aligned} \tag{9}$$

Here $\sigma(t)$, the volatility of the spot interest rate at date t, can depend on all information available at time t. In particular, $\sigma(t)$ can depend on the level of the spot interest rate itself, while $\kappa(x)$ is some deterministic function.

For this choice of the volatility, Ritchken and Sankarasubramanian (1995) show that bond prices at time t can be analytically computed according to

$$P(t,T) = \left(\frac{P(0,T)}{P(0,t)}\right) e^{-\beta(t,T)(r(t)-f(0,t))-\frac{1}{2}\beta^2(t,T)\phi(t)}, \tag{10}$$

with

$$\beta(t,T) = \int_t^T h(t,u)du$$

and

$$\phi(t) = \int_0^t \sigma^2(u)h^2(u,t)du.$$

[5] Notice that most of the interest rate derivatives can be written as portfolios of derivatives on discount bonds.

As before, European options can be computed according to (4) where the expectation is now taken under the risk-neutral process

$$dr(t) = \mu(r,t)dt + \sigma(t)d\tilde{W}(t), \tag{11}$$

with

$$\mu(r,t) = \kappa(t)[f(0,t) - r(t)] + \phi(t) + \frac{d}{dt}f(0,t). \tag{12}$$

In contrast with the process described by equation (7), the process described by equation (11) can be discretised in a Markovian (or recombining) lattice in terms of two variables, namely $r(t)$ and $\phi(t)$. An efficient way to do this has been developed by Li, Ritchken and Sankarasubramanian (1995).

We show now how the models under consideration in this paper (Table 1), can be embedded in the general framework discussed so far. The constant volatility model (HL) can be written in the form of equation (9) by setting $\sigma(t) = \sigma$ and $\kappa(x) = 0$, so that $h(t,T) = 1$. The computation of the functions $\beta(t,T)$ and $\phi(t)$ is trivial. Nesting the exponential volatility model (HW) in (9) is very simple as well. We set $\sigma(t) = \sigma$ and $\kappa(x) = \frac{\lambda}{2}$. This implies that $h(t,T) = \exp[-\frac{\lambda}{2}(T-t)]$. Again, the values for $\beta(t,T)$, $\phi(t)$ immediately follow. Finally we can embed the humped volatility model (MV) in (9) as follows

$$\begin{aligned} \sigma(t) &= \sigma, \\ h(t,T) &= \frac{1+\gamma T}{1+\gamma t} e^{-\frac{\lambda}{2}(T-t)}, \end{aligned}$$

where $\kappa(x) = \frac{\lambda}{2} - \frac{\gamma}{1+\gamma x}$.

The computation of the values of $\beta(t,T)$ and $\phi(t)$ for this model is not as straightforward as before. They are

$$\beta(t,T) = \frac{2}{\lambda^2(\gamma t + 1)}\left[(\gamma\lambda t + 2\gamma + \lambda) - (\gamma\lambda T + 2\gamma + \lambda)e^{-\frac{\lambda}{2}(T-t)}\right],$$

and

$$\begin{aligned} \phi(t) =& \sigma^2 \int_0^t \left(\frac{1+\gamma t}{1+\gamma u}\right)^2 e^{-\lambda(t-u)} du \\ =& \frac{\sigma^2(1+\gamma t)}{\gamma^2}\left[\lambda Ei\left(\frac{\lambda(1+\gamma t)}{\gamma}\right) e^{\frac{-\lambda(1+\gamma t)}{\gamma}}(1+\gamma t) - \gamma\right] \\ &- \frac{\sigma^2(1+\gamma t)^2}{\gamma^2}\left[\lambda Ei\left(\frac{\lambda}{\gamma}\right) e^{\frac{-\lambda(1+\gamma t)}{\gamma}} - \gamma e^{-\lambda t}\right], \end{aligned}$$

where Ei denotes the exponential integral function[6] defined as

$$Ei(z) = \int_{-\infty}^{z} \frac{e^t}{t} dt. \tag{13}$$

[6]Moraleda and Vorst (1996) provide an approximation of this function for any desired accuracy.

As mentioned before, Li, Ritchken and Sankarasubramanian (1995) discretised the spot rate processes of the class identified by Ritchken and Sankarasubramanian (1995). Such a discrete-time model is Markov and hence, the nodes in the lattice recombine. For doing so, they proposed a change of variable which is no longer needed for the models considered in this paper[7]. The discretised lattice approximation for $dr(t)$ in (11) can be established through the following procedure. Suppose that at the start of some time increment the spot rate is r^a. In the next time period the variable moves to either r^{a+} or r^{a-} whose values are given by

$$r^{a+} = r^a + \sigma\left[(J+1)\sqrt{\Delta t}\right],$$
$$r^{a-} = r^a + \sigma\left[(J-1)\sqrt{\Delta t}\right],$$

where J is computed as follows. Set $Z = \text{int}[\frac{\mu(r^a,t)\sqrt{\Delta t}}{\sigma}]$, and $\delta = sign(Z)$. We put $J = \delta Z$ if Z is even and $J = \delta\,|Z|+1$ otherwise. The remaining procedure closely follows the paper by Li, Ritchken and Sankarasubramanian (1995) and we refer to that paper for completeness. That paper also shows how to value American-style options with this recombining lattice.

3 The data

This section describes the two different data sets that are used in this paper. The first data set contains historical yield curves from the Spanish Government bond market. That is, prices of coupon bearing Treasury bonds. However, the term structure of interest rates is not directly observable for most maturities, since the term structure consists of prices of zero coupon bonds. It has to be estimated, basically from bond prices. We devote some attention in this section to detail in which way the term structure has been estimated. The second data set records interest rate option prices. In particular, we took data on options on 10-year Treasury bond futures traded on MEFF (Spanish Exchange of Financial Futures).

3.1 The term structure data

The term structure of interest rates provides a characterization at a specific date of interest rates as a function of time. There are three equivalent ways of specifying this characterization: the discount function, the spot rates for different maturities and the forward rates for different maturities. These functions can be readily computed only when zero coupon bond prices (sometimes referred to as strips) are quoted in the debt markets. However, this is not the case for most markets and certainly not for the Spanish Government bond market. In the most general case, estimation from coupon-bearing bonds has to be implemented to infer the term structure.

[7]In the general form of the Li, Ritchken and Sankarasubramanian (1995) model, stochastic volatilities are possible. To compute such a model, these authors propose a change of variable in order to obtain a constant volatility process. In our cases given by Table 1, this is not required since the instantaneous spot rate volatility is already constant.

In the literature, spline techniques are used to estimate the discount function. McCulloch (1971, 1975) introduced the methodology of fitting the discount function by polynomial splines of different degrees. Shea (1982) and Steely (1991) suggested using B-splines. Vasicek and Fong (1982) implemented exponential splines. Which of these techniques to use for estimating the term structure is not a trivial matter at all. While the estimation should be smooth, it should also approximate as accurate as possible the actual yield curve[8]. It is likely that in the market, options on coupon bearing bonds will be priced based on the market price of the underlying asset. If there is a large difference between the theoretical price from a yield curve model of a coupon bearing bond and its market price, it will be very probable that also the theoretical value of the option is incorrect. Not only will the mispricing in the option be approximately equal to the mispricing in the bond in absolute terms, but in relative terms it will certainly be larger since option prices are much lower than bond prices.

The estimation of the term structure used here was carried out by AFI according to the McCulloch's polynomial (cubic) spline technique for estimating the discount function [9]. This produces estimates of the discount function as a continuous function of time and yields the forward rates to be a smooth function. This estimation was done from daily prices of Spanish Government bonds quoted at the Madrid Stock Exchange. The sample we use lasts for 205 trading days covering the period from July 8, 1994 to May 12,1995. On average, there are 17 daily bond prices. These bonds pay annual coupons. In order to have more short term interest rates from 0 to 1 year, daily prices of REPOS on Treasury bills of the Spanish Government were used too.

3.2 Bond Future Option data

We test the models in this paper with options on Treasury bond futures traded on MEFF (Mercado de Futuros Financieros)[10]. The sample lasts for 94 trading days. It covers the period from January 9, 1995 to May 23, 1995. All prices were taken during the time interval that lasts from 16.00 p.m. to the closing time of the market (17.15 p.m.). In case of several trades of the same option, the one that was closest to 16.00 p.m. has been taken. This is because the yield curve data available in this paper were taken also at 16,00 p.m. The sample contains 1085 options of which 538 are calls being the remaining, 547, puts. On average, there are $12,92$ daily options being traded: $5,83$ calls and $5,81$ puts.

[8]The question regarding which of the above mentioned techniques best approximates the yield curve is not addressed here.

[9]AFI stands for Analistas Financieros Internacionales (International Financial Analysts). We thank Amadeo Reynes and Inmaculada Gomez from AFI that gently provided us with the estimation of the discount function, that, as explained in this section, fully characterize the term structure of interest rates. Thus, we are given the estimated coefficients for the discount function for each day in the sample (for the different intervals into which the time horizon is divided). For doing this estimation, AFI used 5 knots in all the dates considered, i.e., the time horizon was divided in 4 parts. The size of these parts changes daily according to the usual criteria.

[10]Spanish Exchange of Financial Futures. We thank Miguel Angel Rodriguez from MEFF who gently provided us with the option and future data.

The underlying asset of the option is a 10-year Notional bond future traded as well on MEFF. The underlying asset of the future is a notional bond theoretically issued at par the day of the maturity of the future, with a maturity of 10 years, an annual coupon payment of 9% and a face value of 10.000.000 Pts (roughly $ 82.000). There are, basically, four future contracts at any time traded on MEFF, with maturities in March, June, September and December. Specifically, the future contracts expire the third Wednesday of the maturity month. We collected closing prices of Treasury bond futures traded on MEFF. The closing price is computed by MEFF as the average of a number of trades (generally the last 12 trades before the market closes at 17.15 p.m.). The sample period was January 9, 1995 to May 23, 1995, making a total of 94 trading dates. In this period, two bond futures were mainly traded: March 95 and June 95.

The options on these bond futures are American-style. There always trade two different maturities of options. These are the first Wednesdays of the two closest months in which the bond futures mature, i.e., March, June, September and December. An additional option can be traded. It is written on the closest time-to-maturity future and it is referred to as the monthly option. The maturity of this option is also the first Wednesday of the closest month, but just in case that such a month is not March, June, September or December. It means that the monthly option is just called into existence for those months in which no future matures. Eventually, we have then an option expiring the first Wednesday of every month. The underlying asset is always the closest time-to-maturity future. However, there is an important difference between the qualified options which mature in March, June, September or December and the remaining ones. The two closest maturity of the former are always alive while just one -at most- of the monthly options can be traded at any point in time[11].

4 The methodology for the estimation of the volatility functions.

There are two ways for estimating the parameters of the models outlined in the previous sections. One possibility is to use time series of government bond price data to estimate the parameters that specify the dynamics of the underlying interest rate process. The other choice is to estimate implied volatilities from option price data.

The procedure for estimating the implied volatility function is straightforward and very similar to that used for stock options. The problems that might arise are due to the more complicated expressions of interest rate movements and interest rate options compared to stocks.

We focus in this essay on the estimation through time series of the volatility functions from a data set of changes in forward rates. The data was divided into 10 partly overlapping

[11]Recently, the life of the monthly options have been enlarged by MEFF to allow for rolling-over the positions at any time. Thus, either one or two monthly options can always be traded. However, this does not apply to our sample period.

samples of 120 days each. For each trading date in these samples, t_i, the changes in forward rates for maturities T_j, with $T_j \in \{0.5, 1, 1.5, ..., 9.5\}$, were taken. As a result, each basic sample includes 120 observations of 19 variables. Our data, then, take the following form:

$$\Delta f(t_i, T_1)........\Delta f(t_i, T_m) \qquad\qquad i = 1, 2, ..., n \tag{14}$$

where $\Delta f(t_i, T) = f(t_i + 1, T) - f(t_i, T)$ for $i = 1, .., n$ and $T = T_j$ with $T_j \in \{T_1 ... T_m\}$; and $\Delta t_i = t_{i+1} - t_i$. Note that keeping constant the set of forward rate maturities implies that as we move on the sample, t_i, the time-to-maturity of the forward rates, $\tau = T_j - t_i$, decreases.

To estimate the volatility functions of the models, we first of all need to discretise the process (6). Let us define for $h > 0$ the discrete trading interval, where N intervals of size h compose a unit of time. In such a discrete time model, the instantaneous forward rate is defined by:

$$f(t, T) = -\frac{\log(P(t, T + h)/P(t, T))}{h},$$

where $f(t, T)$ denotes the forward interest rate at time t for the investment period $[T, T+h]$ of length h, which is the time interval of the model.

Heath, Jarrow and Morton (1990) have proved that the discrete-time setting of the process (6) is given by

$$\begin{aligned} f(t, T) =& f(0, T) + \sum_{j=1}^{\bar{t}} a_j \sigma(jh, T) C \\ &+ \sum_{j=1}^{\bar{t}} \frac{1}{h} Log \left\{ \frac{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}} \sigma(jh, ih) Ch\right) - 1\right)}{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}-1} \sigma(jh, ih) Ch\right) - 1\right)} \right\}, \end{aligned} \tag{15}$$

where $C = \sqrt{h/(q(1-q))}$ and q stands for objective probabilities,[12] with $0 \leq q \leq 1$; a_j is a Bernoulli random variable, taking the value 1 with probability q and the value 0 with probability $(1-q)$, p is the risk neutral probability

According to (15), the forward rates will move upward if the Bernoulli random variable a takes the value 1, and downward otherwise. The size of the movements are determined by the function σ. Therefore, the variable determining the direction of the movements of the forward rates is a_j in the model. This variable is sometimes referred to as the state variable.

[12]The objective probabilities are those of the general process (3), while the risk adjusted probabilities are those of the the process under the no-arbitrage restriction (5).

For the constant volatility model (HL), the general model in (15) reduces to (HJM 1990a, page 430)

$$f(t,T) = f(0,T) + \sum_{j=1}^{\bar{t}} a_j \sigma C + \frac{1}{h} Ln\left(\frac{1+p(e^{-\overline{T}\sigma C}-1)}{1+p(e^{-(\overline{T}-\bar{t})\sigma C}-1)}\right) \tag{16}$$

with σ being a strictly positive constant. Appendix A shows how the general process (15) simplifies for the HW and MV model.

We then run a principal component analysis. The central idea of principal component analysis is to reduce the dimensionality of a data set in which there is a large number of interrelated variables while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. Specifically, linear combinations of the observed variables are formed. The first component is the combination that accounts for the largest amount of variance in the sample. The second principal component accounts for the next largest amount of the remaining variance and is uncorrelated with the first one. Successive components explain progressively smaller portions of the total sample variance, and all are uncorrelated with each other. The purpose of this technique is to reduce the data while keeping as much information as possible from the original data set. The initial data set is transformed into two matrices, the component or factor scores matrix and a matrix relating these components to the original variables. These later elements are called components or factor loadings (characteristic vectors).

In a number of studies,[13] principal component analysis has been used to determine a reduced number of variables to explain the stochastic evolution of the term structure of interest rates over time. Typically, the historical variability of the rates is almost completely explained with three orthogonal factors. In almost all cases, three factors account for more than 95% of the variability in the data. The first factor essentially represents a parallel shift in the yield curve, while the second and third factors describe the changes in the overall slope and curvature of the yield curve, respectively. These factors (factor loadings) constitute the volatility function of the Heath, Jarrow and Morton model (1990b). In particular, they assume proportional volatility functions given by

$$\sigma(t,T) = \sigma(T-t)f(t,T). \tag{17}$$

Assuming proportional volatility functions as (17) would have allowed us to take the factor loadings from the principal component analysis as the estimators of these functions. This is basically what Heath, Jarrow and Morton (1990b) do leading to a straightforward estimation of the model. Moreover, such volatility functions match market data. But it results in an evolution of the term structure that is non-Markovian with the dramatic reduction in computational tractability that such models imply. Note indeed that the

[13] Dybvig (1989); Heath, Jarrow and Morton (1990b); Steely (1991) and Strickland (1993) are relevant examples.

function (17) cannot be embedded in the Ritchken and Sankarasubramanian (1995) class of volatilities given by (9).

Instead, simpler choices such as those in Table 1 lead to term structure dynamics that can be represented according to Markov processes as shown in section 2. Unfortunately, these specifications of the volatility cannot retain as much explanatory power as the former proportional ones. Furthermore, the volatility functions in Table 1 cannot be directly inferred from a principal component analysis. This is because these specifications of the volatility function no longer directly match the principal component extracted from the historical forward changes[14].

An indirect estimation for the discrete-time processes has been proposed by Hess (1994). He proposes to take the factor scores as estimators of the state variables a_j in (15). As previously outlined, the state variables determine the direction of the movements of the yield curve. Thus if $a_j = 1$, the forward rates move upward, while if $a_j = 0$, the movement of the term structure is downwards. The size of the movements of the yield curve, and how they affect each maturity of the forward rates, is determined by the specification of the volatility function. As a very first approximation, we can proceed as follows. Let the state variable a_j be equal to 1 if the $j-th$ factor score occurs to be positive at time t_j. Consequently, when the factor scores are negative we set a_j equal to 0. But this would just be a rough approximation. Implicitly, we would be assuming that the time interval between observations (Δ) and in the model (h) are the same and, hence, just one movement of the yield curve per trading day would be suitable in our estimation procedure.

To obtain a better fit, we can reduce h, the time step in our model. This would allow for several movements of the yield curve per trading day in the model, though just the daily effect is observed. Then, the factor scores would no longer be estimated realizations of singular Bernoulli random variables. Instead, they would be sums of Δ/h realizations of the state variable. Thus, if we set $\Delta = nh$, then $a_j = \Sigma_{i=1}^{n} s_i$ where s is a Bernoulli random variable taking the values $\{0, 1\}$. The factor scores would be estimators of $a_j \in \{0, 1, 2, .., n\}$, where a_j is the number of upward movements within the period and $(n - a_j)$ notes the number of downward movements.

The problem we face is, inferring the state variables in our model, sums of Bernoulli random variables, from the outcome of the factor analysis (specifically, from the factor scores). We proceed as follows. Let, e.g., $n = 3$, so that $\Delta = 3h$, $a_j = \sum_{i=1}^{3} s_i$ and the factor scores being estimators of $a_j \in \{0, 1, 2, 3\}$, where a_j accounts for the upward movements and $(3 - a_j)$ notes the downward movements. In order to maintain similar probability structures, it is natural to do as follows. The intervals for which the factor scores are higher than 1 are understood as three upward movements in that particular

[14]Heath, Jarrow and Morton (1990b) choose a volatility as (17). They use a data set of proportional changes in forward rates in order to run a principal component analysis. To directly match the outcome of such analysis, they stagger the observation in a convenient way (disregarding some forward rate maturities). In particular, they just look at forward rates whose time-to-maturity is given by $\tau_i = (T_{j+i} - t_j)$ with i denoting the different maturities considered and j standing for the different dates in the sample. By doing so, the outcome of the component analysis exactly matches their volatility function, which was precisely a funtion of the time-to-maturity. See Heath, Jarrow and Morton (1990b)

trading day and we make $a_j = 3$. If the factor scores (fs) are smaller than -1, the assumption of three downward movements is made and we set $a_j = 0$. Consistently, we make $a_j = 2$ if just two upward movement (and one downwards) happened ($0 \leq fs_j < 1$), and finally we let $a_j = 1$ if $-1 \leq fs_j \leq 0$ which would mean that one upwards (and two downwards) movements took place at the time considered[15].

With the state variables a_j so defined, it is now possible to estimate the parameters of the models in Table 1 with a non-linear optimization routine[16]. The function to be minimized is defined as the squared errors of the difference between the forward rate changes given by the models and the observed (historical) forward rate changes. To compute the forward rate changes given by the models we use the previously estimated state variables a_j.

5 Results

For a fixed maturity T, the evolution of the forward rates was modelled by (3). Consider now the set of forward rate observations in our sample given by the maturity set $T_1, ..., T_m$ and arranged as shown in (14). Notice then that T_1 determines the largest data set to be considered: that one where t_n equals T_1.

On the other hand, the largest data set possible would be desirable in order to obtain accurate estimates and to analyse the stability of the parameters through time. However, as explained by Hull (1993), the volatility function changes over time, and data that is too old may no longer be relevant. Furthermore, if, for example, T_1 is chosen to be two years, then the shortest maturity forward rates are disregarded for most of the observations. As a result, T_1 should be chosen in a way that allows for sufficient historical observations, but reasonably small for the short term yield curve not to be systematically disregarded. The choice in this paper is $T_1 = 0.5$, *i.e*, half a year or 120 trading days. Hence, our largest possible choice for t_n is, as well, half a year[17].

The starting date in our option sample is January 9, 1995. At the preceeding trading date, January 6 1995, we take backwards 120 daily yield curve observations. For this data set, we estimate the model parameters as explained in section 4. We can then use the models in Table 1 to price the options in the market. However, as time goes by, new information may arise in the markets. It would not be too realistic then to keep the estimated volatility constant for a long time. It is our choice in this paper to re-estimate the model parameters every two weeks. For doing so, we take backwards again 120 trading

[15]This choice for the intervals of the factor scores seemed natural to us in relating the coefficients a_j and the factor scores from the factor analysis. Notice, in fact, that the probabilities associated with a standard normal distribution for the intervals {(-∞,-1], (-1,0], (0,1], (1,∞)} are {0.15, 0.35, 0.35, 0.15}. In the case of the considered state variables, the probability distribution for a_j= {0, 1, 2, 3} is {1/8, 3/8, 3/8, 1/8}. However, different choices may be possible as long as they maintain a relationship as shown for the coefficients and the factor scores.

[16]We used a Newton Raphson routine that was run in Gauss.

[17]Hull (1993), facing the historical estimation of stock volatility, advises a data set ranging from 90 to 180 daily observations.

days term structure observations and we re-estimate the model parameters. These new estimators are used for pricing options during the immediately following two weeks. After this the whole procedure is repeated. Over all, this routine is implemented ten times. The following subsection details the results of the estimation procedure.

5.1 Estimation of the volatility function

As explained in section 4, we estimate the state variables of the models from a principal component analysis and from this, model parameters are estimated running an error minimization routine. The function to minimize, was defined as the squared errors of the difference between the forward rate changes given by the models and those observed in the markets. The remaining errors will be referred to as residual errors. Clearly, the lower the errors are, the higher the explanatory power of the model. The precise coefficients to be estimated are the volatility function, $\sigma(t,T)$, and the probabilities p and q as given in (15). In all cases p and q have to be estimated whereas the number of volatility parameters varies across the models.

Table 2 shows the parameter estimates for the three models considered for the ten sample periods. However, the estimated values for the objective probability, q, are not tabulated. Their values turned out to be 0.5 for all models and samples considered. This is in agreement with previous studies (see Hess (1994)). Moreover, it could have been to some extent expected, since the forward rate dynamics in (15) are hardly sensitive to the probability parameter p. Therefore, estimated values for p were not expected to differ much from those provided as initial condition in the estimation procedure.[18] The residual error (Res. err.) for each model and sample is also given in Table 2. It follows that the humped volatility model reduces the residual errors most, as could have been expected since this is a three parameter model. We shall come back to this point later in this section.

Table 2 also shows that γ is always positive for the MV model. Humped volatility shapes are just obtained in this model if $2\gamma > \lambda$. It is confirmed in Table 2, that this is the case for all samples considered apart from sample V, in which $2\gamma < \lambda$, so that the volatility is a decreasing function of time. Overall, therefore, nine out of the ten samples studied give empirical support to the humped volatility model. The shape of the hump is not uniform along the samples, though. In fact, the hump is rather sharp for samples I, II, III and IV while it is smoother and wider for samples VI to X. Figure 1 plots some of these shapes for illustrative purposes. Specifically, we draw the volatility function, $\sigma(t,T)$, for the MV model estimated from samples I, III, $VIII$ and X (solid line). Notice that the volatility function for the HW model is also plotted (dotted line).

Comparison of the HW and MV volatility functions in Figure 1 leads to an interesting point. Indeed, λ in the HW model turns out to be negative in samples IV and VI to X making positive the overall sign of the exponential coefficient of this model.[19] Clearly, such an increasing volatility function is implausible for all maturities since it would lead to the

[18] The initial value for p in the estimation routine was chosen to equal 0.5 for all samples and models considered. As explained below, this is its limiting value.

[19] Similar results for λ in the exponential volatility model were found by Amin and Morton (1994).

Table 2: Models parameter estimates

			I					II		
	σ	λ	γ	p	Res. err.	σ	λ	γ	p	Res. err.
HL	0.0146			0.5171	0.008548	0.0142			0.4781	0.015835
HW	0.0150	0.0108		0.5072	0.008543	0.0158	0.0306		0.4790	0.015602
MV	0.0047	0.4587	2.4401	0.5001	0.007024	0.0042	0.5244	3.9045	0.4760	0.009352
			III					IV		
	σ	λ	γ	p	Res. err.	σ	λ	γ	p	Res. err.
HL	0.0134			0.4737	0.014167	0.0146			0.5102	0.021490
HW	0.0161	0.0613		0.4752	0.012827	0.0076	-0.1501		0.4989	0.020864
MV	0.0060	0.5162	1.7886	0.4716	0.006459	0.0026	0.2208	2.3531	0.4760	0.017835
			V^*					VI		
	σ	λ	γ	p	Res. err.	σ	λ	γ	p	Res. err.
HL	0.0117			0.4964	0.003562	0.0157			0.4889	0.006677
HW	0.0119	0.0070		0.4966	0.003561	0.0155	-0.0042		0.4887	0.006676
MV	0.0141	0.0844	0.0052	0.4968	0.003476	0.0112	0.0827	0.1517	0.4884	0.006379
			VII					VIII		
	σ	λ	γ	p	q	σ	λ	γ	p	Res. err.
HL	0.0148			0.4846	0.008671	0.0110			0.4667	0.009941
HW	0.0116	-0.0932		0.4799	0.007908	0.0086	-0.0751		0.4549	0.008622
MV	0.0085	0.0066	0.1460	0.4834	0.006902	0.0061	0.2232	0.4651	0.4884	0.007783
			IX					X		
	σ	λ	γ	p	Res. err.	σ	λ	γ	p	Res. err.
HL	0.0110			0.4758	0.010543	0.0183			0.5156	0.008639
HW	0.0085	-0.1452		0.4722	0.007789	0.0143	-0.0844		0.5148	0.007793
MV	0.0039	0.2464	1.3158	0.4759	0.007480	0.0112	0.1851	0.3680	0.4884	0.007489

Roman numbers stand for the 10 samples whose period are reported below. Each of these samples account for 120 daily observations of the yield curve. Sample I: July 8 1994 to January 6 1995; Sample II: July 22 1994 to January 20 1995; Sample III: August 8 1994 to February 3 1995; Sample IV: August 23 1994 to February 17 1995; Sample V: September 6 1994 to March 3 1995; Sample VI: September 21 1994 to March 17 1995; Sample VII: October 4 1994 to March 31 1995; Sample VIII: October 17 1994 to April 12 1995; Sample IX: October 31 1994 to April 28 1995; Sample X: November 14 1994 to May 12 1995.
(*) All parameters reported for sample V are estimated from the sample period September 21 1994 to March 3 1995. This is a subsample accounting for 100 days rather than 120. The estimation for the sample September 6 1994 to March 3 1995 turns out to be highly biased while that one reported seems steadier.

explosion of interest rates. Therefore, negative values for λ in the exponential volatility model should be precluded. But this reduces the model to the constant volatility one for samples IV and VI to X.[20] Notice moreover, that quite apart from the theoretical disadvantages of the constant volatility model, the differences in the residual errors become now apparent. This can be verified by comparing the residual errors (Res. err.) in Table 2 for the HW and MV model in samples I, II, III and V; and the HL and MV model in samples IV and VI to X where, as explained, the HW model reduces to the HL model.

It should be noticed however, that the estimate for λ in the HW model results to be positive for samples I to III even though humped volatility structures are found for these samples with the MV model. As Figure 1 illustrates, this is because the downward part of the hump is dominant.[21] Although Figure 1 sufficiently clarifies this fact, it can be further verified by comparing the residual errors of both models in Table 2 for samples I to III.

In conclusion, not only serious inconsistencies are found when estimating the exponential volatility model, but the fit to real yield curve data is largely improved by using the humped volatility model. Our yield curve data, frequently reveals that the mean reverting effect is not as straightforward as generally believed.

We address next the analysis of the results for the probability parameters in Table 2. This table shows that both the risk neutral and objective probabilities, p and q respectively, are very close to 0.5. In particular, q is systematically equal to 0.5 whereas p always lies in the interval $[0.45, 0.52]$. Moreover, the forward rates are hardly sensitive to changes in the objective probability coefficient, q. In turn, the yield curve dynamics are very sensitive to changes in the risk neutral probability p. This proximity to 0.5 of both parameters is in agreement with a theoretical result by Heath, Jarrow and Morton (1990a). They proved that the discrete-time process (15) converges in probability to its continuous-time counterpart process (6) if and only if $p = q = 0.5$ in (15). Thus, the limiting forward rate process is insensitive to the parameters p and q. Accordingly, these authors argue that an estimation procedure that involves inverting the contingent claims values (implied volatility estimation from option data) should be done with the limiting processes. This is because the option market is a continuous-time economy and, hence, the option value should be insensitive to the probabilities (see Heath, Jarrow and Morton (1990a), page 438). Whether this also applies to our model is discussed next.

In particular, we want to study the estimates of the volatility function achieved by the discrete-time process where no restrictions are imposed on p and q, and the limiting process where p and q are forced to take their limiting value 0.5. We study four subsamples of each of the ten samples reported in Table 2. More precisely, the last date of each sample we take backwards 120 (entire sample whose results are shown in Table 2), 100, 80 and 60 observations. These subsamples are referred to as A, B, C and D, respectively. For ease of exposition we restrict our attention to the constant volatility model (HL).

[20] The volatility coefficients as shown in Table 2 have been used for pricing bond future options whose results are studied in the next section. This is similar to what Amin and Morton (1994) do, but as explained, if the model is to be used practically, negative values for λ should be precluded.

[21] Note that the volatility functions in Figure 1 are plotted for 25 years to maturity. However we just use yield curve data up to 10 years for estimating model parameters.

Figure 1: Volatility function for the HW (dotted line) and MV (solid line) model

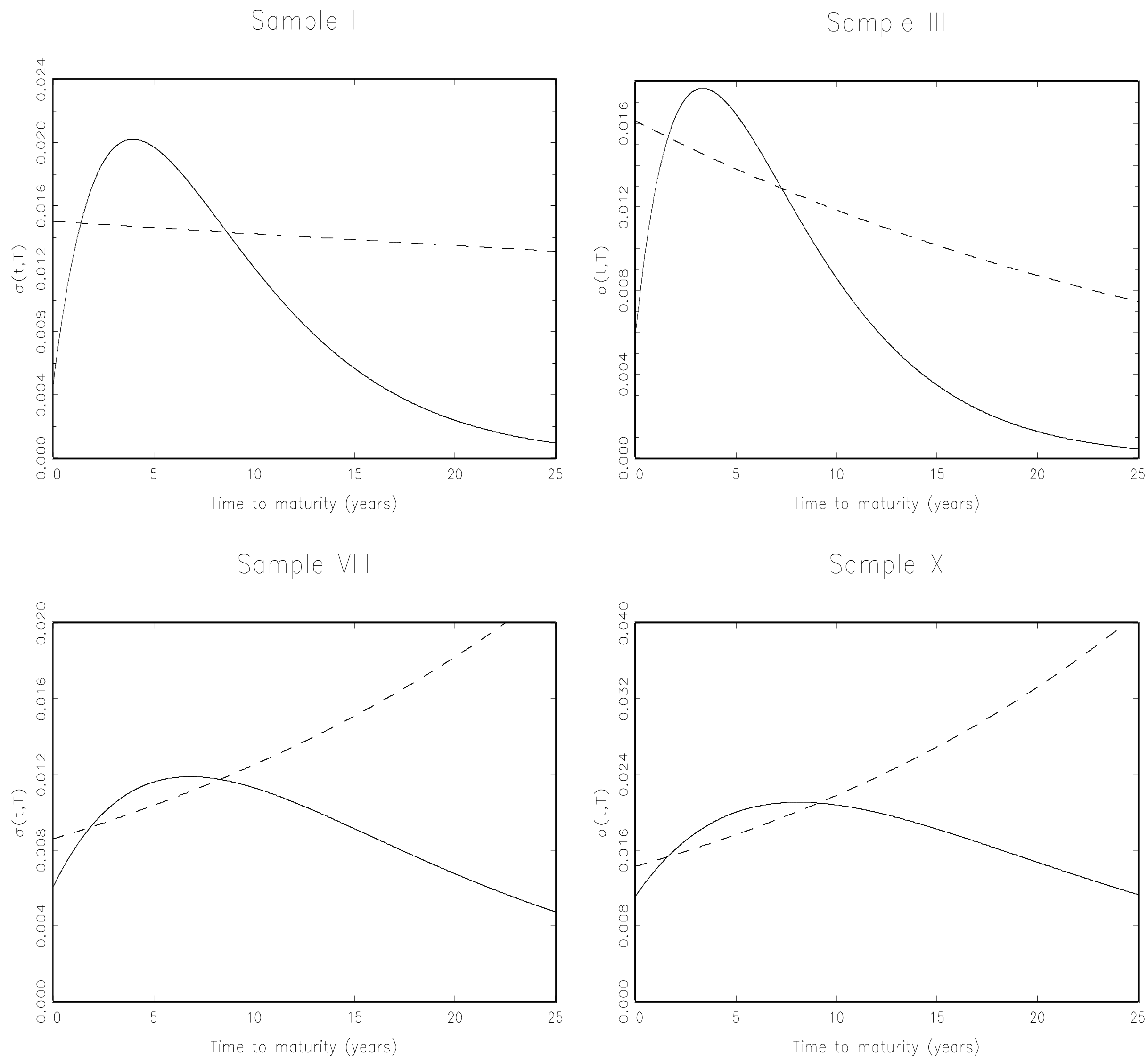

Table 3: Estimates of σ for the constant volatility (HL) model

	I		II		III		IV		V	
	dt	ct	dt	ct	dt	ct	dt	ct	dt	ct
A	0.0146	0.0106	0.0142	0.0290	0.0134	0.0251	0.0146	0.0081	0.0171	0.0110
B	0.0124	0.0190	0.0113	0.0148	0.0158	0.0059	0.0137	0.0047	0.0117	0.0125
C	0.0134	0.0121	0.0147	0.0100	0.0127	0.0085	0.0139	0.0152	0.0109	0.0040
D	0.0129	0.0089	0.0135	0.0037	0.0119	0.0136	0.0144	0.0163	0.0163	0.0211

	VI		VII		VIII		IX		X	
	dt	ct	dt	ct	dt	ct	dt	ct	dt	ct
A	0.0157	0.0082	0.0148	0.0170	0.0110	0.0285	0.0110	0.0231	0.0183	0.0101
B	0.0140	0	0.0111	0	0.0102	0	0.0152	0.0352	0.0172	0.0105
C	0.0111	0	0.0144	0.0053	0.0166	0.0297	0.0165	0.0143	0.0192	0.0042
D	0.0157	0	0.0159	0.0181	0.0219	0.0149	0.0188	0.0184	0.0172	0.0095

Roman numbers, I to X, stand for the sample periods reported in the caption of Table 2. Each of these samples accounts for 120 daily observations. A, B, C, D are subsamples of each of the previous samples. A accounts for the total sample, *i.e.*, 120 observations, or alternatively six month yield curve data. B, C and D correspond to subsamples of 100, 80, and 60 observation respectively. Finally, *dt* denotes the discrete-time processes as given by (15) while *ct* denotes the limiting process, i.e., the probabilities p and q in (15) are fixed equal to 0.5.

Columns *dt* in Table 3 report the volatility parameters for the HL model estimated from the discrete-time process (16) where p and q were allowed to take any value in the interval $[0, 1]$. Columns *ct* in Table 3 shows the estimation of the volatility function using the limiting process for the HL model. As previously explained, the limiting process is obtained by setting $p = q = 0.5$ in (16). Comparing columns *dt* and *ct* in Table 3, it is obvious that the *dt* estimators are more stable than those produced by the limiting process. Moreover, in many cases the *ct* volatility parameters are highly misestimated, yielding in some cases null values for the volatility in the HL model. That is, the process (16) is very sensitive to the risk neutral probability p, and although its value is very close to 0.5 (see Table 2), forcing it to take this exact value leads, in many cases, to a misspecification of the volatility coefficients. As a consequence, the probability parameters should be left free.

5.2 Yield curve and future prices

Pricing options with the models in Table 1 just requires as inputs the initial yield curve on the valuation date and the volatility function parameters.[22] The volatility function is estimated as shown in section 4. The initial yield curve has to be estimated from bond prices as explained in section 3. It was explained in section 3 that the estimation of the term structure of interest rates is not trivial. Moreover, such estimation faces a tradeoff between smoothness and accuracy. To give some insights on the accuracy of

[22]This is, of course, apart from the specific features of the options to be priced.

Table 4: Summary of bond future pricing errors

Bond Future	Av. err.[a]	St. dev.[b]	Av. rel. abs. err.[c]
March-95	-0.35	0.23	0.0044
June-95	-0.60	0.43	0.0080

The sample period is January 6, 1995 to May 26, 1995. The total number of trading dates in the sample is 94. During 36 trading dates options with the March-95 bond future as the underlying asset were being traded. The June-95 bond future was the underlying asset of traded options on 91 trading dates. The price of the bond future is the same for all models considered. It just depends on the initial yield curve. (a) Av. err. = Average error where the pricing error is defined as model price minus market price. (b) St. dev. err. = Standard Deviation of the errors defined as (a).(c) Av. rel. abs. err. = average of (Pricing error/Market price).

the estimation of the term structure, we compute the errors we make in pricing bond futures. The 10 years Treasury bond futures are traded in MEFF, so that market prices are available. Their model price can be calculated from the yield curve. Moreover, these are the underlying instruments of the options we consider later in this chapter. Table 4 reports some descriptive statistics of the pricing errors defined as model price minus market price. Notice that the models systematically underprice the futures in the market. The errors though are not large. On average, we make an error of -0.35 and -0.60 percentage points over the nominal value of the bond futures of March 95 and June 95, respectively. This means a relative absolute error of, on average, 0.0044 and 0.0080.[23] It is important to realize that these errors may partially be caused by a lack of synchronicity of the data. Recall that the yield curve data was estimated with bond prices taken at 16.00 p.m. while the bond future data consists of closing prices. The closing price is computed by MEFF as the average of a number of trades (generally the last 12 trades before the market closes at 17.15 p.m.). Due to the liquidity of these instruments, most of the references for calculating the closing price are taken in the interval from 17.00 to 17.15. Therefore, there is roughly a gap of one hour between the different data sets. Given the underpricing of futures, it is likely that all option pricing models produce biased theoretical values. However, all models price bond futures exactly in the same way. They will all be affected by the same bias and, hence, the comparison between the models is still possible.

5.3 Option pricing

Descriptive statistics for pricing errors of options produced by the different models are shown in Table 5. The pricing error is defined as model price minus market price. All

[23]Both market and model prices of the bond futures are given in percentage over its nominal value. Thus, the average and standard deviation of the pricing errors are also given in the same unit. In turn, the average relative absolute error reported in the last column of Table 4 is computed as pricing error over market price.

Table 5: Descriptive statistics for pricing errors

All Options					
Model	Av. error[a]	Var. err[b]	Av. abs. err[c]	Av. rel. abs. err.[d]	Sum sqr.err.[d]
HL	-0.09	0.09	0.21	0.39	103.42
HW	-0.11	0.09	0.22	0.40	106.19
MV	-0.11	0.10	0.24	0.45	118.93
Call Options					
Model	Av. error[a]	Var. err[b]	Av. abs. err[c]	Av. rel. abs. err.[d]	Sum sqr.err.[d]
HL	-0.17	0.07	0.22	0.40	57.31
HW	-0.20	0.07	0.25	0.45	62.20
MV	-0.20	0.08	0.26	0.47	67.24
Put Options					
Model	Av. error[a]	Var. err[b]	Av. abs. err[c]	Av. rel. abs. err.[d]	Sum sqr.err.[d]
HL	0.00	0.09	0.20	0.37	46.11
HW	-0.02	0.08	0.20	0.36	43.99
MV	-0.02	0.10	0.23	0.42	51.69

(a) Av. err. = Average error where the pricing error is defined as model price minus market price. (b) Var. err. = Variance of the errors defined as (a). (c) Av. abs. err.= average pricing absolute error. (d) Av. rel. abs. err. = average of (Pricing error/Market price). (e) Sum sqr. err. = Sum of the squared pricing error. All prices of both market and models were computed as percentage of the face value of the underlying asset.

option prices are computed as percentage of the face value of the underlying asset. Notice, that the option prices given by the models will be affected by the mispricing in the futures reported in Table 4. In particular, the mispricing in the option will be approximately equal to the mispricing in the bond future in absolute terms, but in relative terms it will certainly be larger since option prices are much lower than bond prices.

Table 5 shows that the constant volatility (HL) model provides the best fit to the data. In fact, it is the model that makes the lowest error on average and the total squared errors are also smaller. However, the mispricing for all models is very similar. They all drastically underprice the call options, while, on average, provide a better fit to the puts.

The main conclusion, however, that can be drawn from Table 5 is that the models poorly match the data. In fact, the average mispricing of the models ranges from 39% at best to 45% in the worst case. In both cases, this is the mispricing percentage over the options quoted in the market.

There are three reasons for these disappointing results. First, we use historical volatility, where options in the markets are priced based on expectation of the volatility of the underlying asset over the remaining time to maturity of the option. If for example, the market expects certain announcements about the interest rate policy of the Central Bank or some other government agency, this might cause more uncertainty. Hence, traders might increase the volatility that they put in their models above the historically observed level. Second, although the derived theoretical prices are arbitrage free prices this of course only holds if the dynamics of bond prices are indeed correctly described by the models. Since, we cannot fully explain observed bond price dynamics with our models the theoretical prices do not have to be arbitrage free prices. It is well known that market participants often use a model developed by Black (1976) for pricing bond options. If one applies this methodology to a large portfolio of options and bonds, Black's model certainly leads to inconsistencies. However recently, Miltersen, Sandmann and Sondermann (1995) have proved that for individual bond options one can specify a dynamic process for the price of the underlying bond such that Black's model gives exactly the no arbitrage price. The inconsistencies are due to the fact that if these processes really describe the price behavior of a portfolio of bonds then there are arbitrage opportunities within the bond portfolio. Finally, the no arbitrage argument is based upon the assumptions of continuous trading and frictionless markets. Hence, small arbitrage opportunities can not be exploited due to the market frictions. Market makers might increase the prices of options if there is a net demand for certain kind of options to just below the level at which arbitrage opportunities can be exploited in markets with frictions. We think that the first reason might be the most relevant for this study.

6 Conclusions

We have studied in this paper the empirical behavior of three one-factor interest rate models as developed by Ho and Lee (1986), Hull and White (1990) and Moraleda and Vorst (1996). The same general setting has been used for all models. In particular, they all have been embedded in the Ritchken and Sankarasubramanian (1995) framework. This is a general model within the Heath, Jarrow and Morton (1992) paradigm, that allows for pricing American options through recombining binomial trees. Moreover, the same estimation procedure has been run in all cases. Specifically, we have used an indirect estimation procedure from a principal component analysis. The model parameters have been estimated from daily yield curve data. The parameter estimates have been used to price bond future options traded in MEFF.

Humped volatility shapes as modeled by Moraleda and Vorst (1996), have been found in 90% of the samples studied. Moreover, serious inconsistencies are found when estimating the exponential volatility model. In particular, the exponential coefficient of this model, λ, has a negative estimated value in 60% of the samples. Since this makes interest rates explode, λ should be precluded to become negative.

The humped volatility model (MV) turns out to be the model that best explains the yield curve movements. In fact, it is the model that most reduces the residual errors but it also has one more parameter. This higher explanatory power is even more evident when restricting the exponential coefficient, λ, of the HW model to be non-negative. However, this ability no longer applies when pricing options. In fact, the simpler constant volatility model (HL) minimizes the option pricing errors. Nevertheless, the errors are rather large for all models. There have been given three reasons that might cause this frustrating result. First, historical volatility might not be a correct indication of expected future volatility in the market. Second, the arbitrage free prices are based on the assumption that the underlying bond prices are correctly described by our models. Finally, market imperfections do not allow the exploitation of all arbitrage opportunities.

Appendix

A Discrete-time models

The discrete time version of the exponential model (HW) is

$$\begin{aligned} f(t,T) =& f(0,T) + \sum_{j=1}^{\bar{t}} a_j \sigma e^{-(\lambda/2)(\overline{T}-j)h} C \\ &+ \frac{1}{h}\sum_{j=1}^{\bar{t}} Log \left\{ \frac{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}} \sigma \exp(-(\lambda/2)(i-j)h)Ch\right) - 1\right)}{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}-1} \sigma \exp(-(\lambda/2)(i-j)h)Ch\right) - 1\right)} \right\} \end{aligned} \tag{A.1}$$

where σ and λ are positive constant.

The humped volatility discrte-time model (MV) is given by

$$\begin{aligned} f(t,T) =& f(0,T) + \sum_{j=1}^{\bar{t}} a_j \sigma \frac{1+\gamma T}{1+\gamma j h} e^{-(\lambda/2)(\overline{T}-j)h} C \\ &+ \frac{1}{h}\sum_{j=1}^{\bar{t}} Log \left\{ \frac{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}} \sigma \frac{1+\gamma i h}{1+\gamma j h} \exp(-(\lambda/2)(i-j)h)Ch\right) - 1\right)}{1 + p\left(\exp\left(-\sum_{i=j}^{\overline{T}-1} \sigma \frac{1+\gamma i h}{1+\gamma j h} \exp(-(\lambda/2)(i-j)h)Ch\right) - 1\right)} \right\} \end{aligned} \tag{A.2}$$

where σ γ and λ are positive constant.

References

[1] K.I. Amin and A.J. Morton (1994): "Implied volatility functions in arbitrage-free term structure models", Journal of Financial Economics, 35, 141-180.

[2] F. Black (1976): "The pricing of commodity contracts", Journal of Financial Economics, 3, 167-179.

[3] F. Black and P. Karasinski (1991): "Bond and option pricing when short rates are lognormal", Financial Analysts Journal, 47, 52-59.

[4] M. Brennan and E. Schwartz (1982): "An equilibrium model of bond pricing and a test of market efficiency", Journal of Financial and Quantitative Analysis, 17, 301-329.

[5] R.H. Brown and P.H. Dybvig (1986): "The empirical implications of the Cox-Ingersoll-Ross theory of the term structure of interest rates", Journal of Finance, 41, 617-630.

[6] R.H. Brown and S.M. Schaefer (1994): "The term structure of real interest rates and the Cox, Ingersoll and Ross model", Journal of Financial Economics, 35, 3-42.

[7] K.C. Chan, G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992): "An empirical comparison of alternative models of the short-term interest rate", Journal of Finance, 47, 1209-1227.

[8] J.C. Cox, J.E. Ingersoll and S.A. Ross (1985): "A theory of the term structure of interest rates", Econometrica, 53, 385-408.

[9] B. Dietrich-Campbell and E. Schwartz (1986): "Valuing debt options, empirical evidence", Journal of Financial Economics, 16, 321-343.

[10] P.H. Dybvig (1989): "Bond and option pricing based on the current term structure", Working paper, Washington University, St. Louis, Missouri.

[11] B. Flesaker (1993): "Testing the Heath-Jarrow-Morton/Ho-Lee model of interest rate contingent claims pricing". Journal of Financial and Quantitative Analysis, 28, 483-495.

[12] H.G. Fong and O.A. Vasicek (1982):"Term structure modeling using exponential splines". Journal of Finance, 37, 339-348.

[13] D. Heath, R. Jarrow and A. Morton (1990a): "Bond pricing and the term structure of interest rates : a discrete time approximation". Journal of Financial and Quantitative Analysis, 25, 419-440.

[14] D. Heath, R. Jarrow and A. Morton (1990b): "Contingent claims valuation with a random evolution of interest rates". Review of Futures Markets, 9, 55-76.

[15] D. Heath, R. Jarrow and A. Morton (1992): "Bond pricing and the term structure of interest rates : a new methodology for contingent claims valuation". Econometrica, 60, 77-105.

[16] D. Heath, R. Jarrow, A.J. Morton, and M. Spindel, (1992): "Easier done than said", RISK, 5, 77-80.

[17] S.D. Hess (1994): "Estimation of discrete time multiple factor models of the term structure of interest rates". University of Konstanz, Faculty of Economics and Statistics, Technical report.

[18] T.S.Y. Ho and S-B. Lee (1986): "Term structure movements and pricing interest rate contingent claims". Journal of Finance, 41, 1011-1029.

[19] J. Hull (1995): "Options, Futures and other derivative securities", third edition, Prentice-Hall.

[20] J. Hull and A. White (1990): "Pricing interest rate derivative securities", Review of Financial Studies, 3, 573-592.

[21] J. Hull and A. White (1993): "One-factor interest rate models and the valuation of interest-rate derivative securities", Journal of Financial and Quantitative Analysis, 28, 235-254.

[22] R. Kahn, (1991): "Fixed income risk modeling". The Handbook of Fixed Income Securities, Third edition, edited by F. Fabozzi, 1307-1319.

[23] A. Li, P. Ritchken and L. Sankarasubramanian (1995): "Lattice models for pricing American interest rate claims", Journal of Finance, vol 1, 719-737.

[24] J.H. McCulloch (1971) :"Measuring the term structure of interest rate". Journal of Business, 46, 19-31.

[25] J.H. McCulloch (1975) : "An estimate of the liquidity premium". Journal of Political Economy, 83, 95-119.

[26] F. Mercurio and J.M. Moraleda (1996): "An analytically tractable interest rate model with humped volatility", discussion paper TI 96-116-2, Tinbergen Institute, Erasmus University Rotterdam, The Netherlands.

[27] K.R. Miltersen, K. Sandmann and D. Sondermann (1995): "Closed form solutions for term structure derivatives with log-normal interest rates", forthcoming in Journal of Finance.

[28] J.M. Moraleda and A.C.F. Vorst (1996): "Pricing American interest rate claims with humped volatility models", discussion paper TI 96-115-2, Tinbergen Institute, Erasmus University Rotterdam, The Netherlands.

[29] P. Ritchken and L. Sankarasubramanian (1995): "Volatility structures of forward rates and the dynamics of the term structure", Mathematical Finance, 5, 55-72.

[30] S.M. Stambaugh (1988): "The information in forward rates: implications for models of the term structure". Journal of Financial Economics, 21, 41-70.

[31] J. Steeley (1991): "Estimating the gilt-edged term structure: basic splines and confidence intervals", Journal of Business Finance and Accounting, 18, 513-529.

[32] C. Strickland (1993): "An analysis of the Heath, Jarrow and Morton approach to pricing interest rate contingent claims", working paper, FORC, University of Warwick

[33] O.A. Vasicek (1977): "An equilibrium characterization of the term structure". Journal of Financial Economics, 5, 177-188.